

## SOEPpapers

on Multidisciplinary Panel Data Research

SOEP – The German Socio-Economic Panel Study at DIW Berlin

407-2011

# Bestimmung der Herkunftsnationen von Teilnehmern des Sozio- oekonomischen Panels (SOEP) mit Migrationshintergrund

Friedrich Scheller

## **SOEPPapers on Multidisciplinary Panel Data Research** at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel Study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPPapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPPapers are available at  
**<http://www.diw.de/soeppapers>**

### **Editors:**

Joachim R. **Frick** (Empirical Economics)  
Jürgen **Schupp** (Sociology, Vice Dean DIW Graduate Center)  
Gert G. **Wagner** (Social Sciences)

Conchita **D'Ambrosio** (Public Economics)  
Denis **Gerstorff** (Psychology, DIW Research Professor)  
Elke **Holst** (Gender Studies)  
Frauke **Kreuter** (Survey Methodology, DIW Research Professor)  
Martin **Kroh** (Political Science and Survey Methodology)  
Frieder R. **Lang** (Psychology, DIW Research Professor)  
Henning **Lohmann** (Sociology, DIW Research Professor)  
Jörg-Peter **Schräpler** (Survey Methodology, DIW Research Professor)  
Thomas **Siedler** (Empirical Economics, DIW Graduate Center)  
C. Katharina **Spieß** (Empirical Economics and Educational Science)

ISSN: 1864-6689 (online)

German Socio-Economic Panel Study (SOEP)  
DIW Berlin  
Mohrenstrasse 58  
10117 Berlin, Germany

Contact: Uta Rahmann | [soeppapers@diw.de](mailto:soeppapers@diw.de)

**Bestimmung der Herkunftsnationen von Teilnehmern des  
Sozio-oekonomischen Panels (SOEP) mit  
Migrationshintergrund**

Friedrich Scheller\*

Duisburg, Juni 2011

\* Universität Duisburg-Essen, Institut Arbeit und Qualifikation, E-Mail: [friedrich.scheller@uni-due.de](mailto:friedrich.scheller@uni-due.de)

## **Abstract**

Die Heterogenität der zugewanderten Bevölkerung in Deutschland spielt bei der Analyse migrations- und integrationsbezogener Fragen eine wachsende Rolle. Seit der Welle Y ist im Sozio-oekonomischen Panel (SOEP) die Variable MIGBACK enthalten, welche die Identifikation des Migrationsstatus der Befragten ermöglicht. Empirisch werden, z. B. bei der Arbeitsmarktintegration, zusätzlich zu Unterschieden zwischen zugewanderter und autochthoner Bevölkerung sowie zwischen den Generationen jedoch immer wieder auch innerhalb der Gruppe der Zuwanderer und ihrer Nachkommen Ungleichheiten entlang der Herkunftsnationen festgestellt. Um diesbezügliche Untersuchungen mit dem SOEP zu vereinfachen, wird gezeigt, wie vorhandene Informationen auch für eine möglichst breite Zuordnung der Herkunftsnationen der Teilnehmer mit Migrationshintergrund verwendet werden können. Durch die Nutzung von Angaben zu früheren Nationalitäten, dem Geburtsland oder der Herkunft der Eltern kann neben der unproblematischen Zuordnung des Herkunftslandes selbst Zugewanderter auch bei einem Großteil der in Deutschland Geborenen mit indirektem Migrationshintergrund die Herkunft identifiziert werden. Die gegenüber der einfachen Bestimmung anhand der Nationalität gewonnenen Fallzahlen werden abschließend nach Herkunftsnationen differenzierend dargestellt und auf bestehende (Fallzahlen-)Probleme und Grenzen der skizzierten Vorgehensweise aufmerksam gemacht.

**Schlagwörter:** Migration, Methoden

## **Inhaltsverzeichnis**

<b>1</b>	<b>Einleitung .....</b>	<b>1</b>
<b>2</b>	<b>Daten .....</b>	<b>2</b>
<b>3</b>	<b>Herkunftsbestimmung: Vorgehensweise.....</b>	<b>3</b>
<b>4</b>	<b>Migrantengruppen mit bestimmter / nicht bestimmter Herkunft.....</b>	<b>8</b>
<b>5</b>	<b>Verteilung nach Herkunftsländern.....</b>	<b>11</b>
<b>6</b>	<b>Schluss .....</b>	<b>15</b>



# 1 Einleitung<sup>1</sup>

Die Heterogenität der zugewanderten Bevölkerung in Deutschland spielt bei der Analyse migrations- und integrationsbezogener Fragen eine wachsende Rolle. Neben der getrennten Betrachtung von Ausländern der ersten und der bereits in Deutschland geborenen nachfolgenden Generationen mit ausländischer Staatsangehörigkeit ist hierbei insbesondere der Migrationshintergrund von zentraler Bedeutung. Die Abkehr von der alleinigen Unterscheidung anhand einer vorhandenen oder nicht vorhandenen deutschen Nationalität ermöglicht einen differenzierteren Blick auf migrationsspezifische Zusammenhänge (vgl. z. B. Diefenbach / Weiß 2006; Gresch / Kristen 2011; Haug 2011). Jedoch kann auch bei einer derartigen Differenzierung nicht zwingend davon ausgegangen werden, dass untersuchte Erklärungsmuster für Migranten unterschiedlicher Herkunftsregionen und -kontexte gleichermaßen relevant sind.

So werden empirisch, beispielsweise bei der Arbeitsmarktintegration, zusätzlich zu Unterschieden zwischen zugewanderter und autochthoner Bevölkerung sowie zwischen den Generationen innerhalb der Gruppe der Zuwanderer und ihrer Nachkommen immer wieder auch Ungleichheiten entlang der Herkunftsnationen festgestellt. Es liegt daher nahe, bei Analysen der Situation von Zuwanderern die Heterogenität der Gruppe der Migranten über den Migrationshintergrund hinaus zu berücksichtigen. Obschon teils gerechtfertigte Einwände zur Aussagekraft von Länderzugehörigkeiten bestehen, ist das Herkunftsland dabei ein relativ eindeutiges Unterscheidungsmerkmal. Dieses muss Unterschiede zwar nicht erklären, kann jedoch bestehende gruppenspezifische Besonderheiten verdeutlichen. Derartige Besonderheiten können bezüglich der sozioökonomischen Zusammensetzung verschiedener Herkunftsgruppen bestehen, aber z. B. auch in der räumlichen Verteilung, der Größe und Nutzung sozialer Netzwerke (vgl. z. B. Granato 2009; Haug 2005; Kalter 2007) oder der ethnospezifisch unterschiedlichen Bedeutung transnationaler Räume (siehe Berger / Weiß 2008; Pries 2008) liegen.

Bei Migranten der ersten Zuwanderergeneration stellt die Zuordnung zu einem Herkunftsland in der Regel kein Problem dar. Soweit diese Informationen vorliegen, kann die Herkunft hier über das Geburtsland bzw. das Land, aus dem die Auswanderung stattgefunden hat oder auch über die Nationalität bestimmt werden. Doch ist häufig insbesondere auch die Situation von

---

<sup>1</sup> Ich bedanke mich bei Marcel Erlinghagen und Belit Şaka für wertvolle Hinweise und Ergänzungen, die sehr zur Entstehung dieses Textes beigetragen haben.

Migranten der zweiten und nachfolgenden Generationen von Interesse. Hier auftretende gruppenspezifische Besonderheiten weisen auf dauerhafte Differenzen einzelner Herkunftsgruppen hin, welche über mögliche Schwierigkeiten in Verbindung mit unterschiedlichen Startbedingungen zum Zeitpunkt der Zuwanderung hinausgehen. Allerdings ist bei den nachfolgenden Generationen die Zuordnung von Herkunftsländern oft weniger eindeutig. Angaben über Geburtsland oder Staatsangehörigkeit reichen bei in Deutschland geborenen deutschen Staatsangehörigen nicht aus. Hier wird es notwendig, weitere Merkmale wie den familiären Hintergrund einzubeziehen. Wie dies mittels der Daten des Sozio-oekonomischen Panels (SOEP) möglich ist, soll hier näher dargestellt werden.

## **2        Daten**

Im SOEP ist seit der Welle Y im Jahr 2008 in den vom DIW veröffentlichten Daten bereits die Variable MIGBACK enthalten, welche die Identifikation des Migrationsstatus der Befragten ermöglicht (vgl. Groh-Samberg et al. 2010). Diese Variable erlaubt die Differenzierung von Personen mit und ohne Migrationshintergrund über die einfache Trennung nach Staatsangehörigkeit hinaus. Die genannten Informationen wurden dabei rückwirkend für sämtliche Wellen zusammengestellt, auch wenn die betreffenden Personen zuletzt nicht teilgenommen haben. Bei einem Großteil der Betroffenen kann zwischen direktem (d. h., die Person ist selbst zugewandert) und indirektem Migrationshintergrund (Person ist in Deutschland geboren) unterschieden werden. Diese weitere Differenzierung ist für über 90 Prozent der identifizierten Befragten mit Migrationshintergrund möglich. Für alle übrigen kann ohne erneute Befragung aufgrund fehlender Informationen zum Geburtsland anhand der im SOEP vorhandenen Variablen der genaue Migrationsstatus nicht festgestellt werden. Dessen ungeachtet stellt die Erweiterung des Datensatzes um den Migrationsstatus eine äußerst hilfreiche Erweiterung für migrationsbezogene Analysen dar. Konnten Herkunftsmerkmale, welche über die Unterscheidung Deutsche vs. Ausländer hinausgehen, zwar zuvor auch eigenhändig von den Datennutzern bestimmt werden (vgl. Frick/Söhn 2005), besteht nun eine, mehr oder weniger, vereinheitlichte und in der Handhabung deutlich vereinfachte Möglichkeit für derartige Analysen.

Wenn aber Untersuchungen von Differenzen zwischen Zuwanderern verschiedener Herkunftsnationen ebenfalls Eingebürgerte und Nachkommen mit deutscher Staatsangehörigkeit einschließen sollen, ist es notwendig, auch hier über die alleinige Berücksichtigung der aktuellen Staatsangehörigkeit hinaus Merkmale wie eine frühere Nationalität, das Geburtsland



oder die Herkunft der Eltern mit einzubeziehen. Das SOEP enthält als Panelbefragung mit von Beginn an starker Fokussierung der Integrationsthematik eine Vielzahl von Variablen, die Rückschlüsse auf das Herkunftsland bzw. die Herkunftsländer der zugewanderten Vorfahren zulassen, auch wenn diesbezügliche Informationen in der jeweils aktuellen Befragungswelle nicht erfragt oder hierzu keine Angaben gemacht wurden. Als Haushaltsbefragung bietet es zudem die Möglichkeit, auf eventuelle Angaben der Eltern zurückzugreifen. Diese Informationen wurden bislang (soweit bekannt) nicht gebündelt zusammengetragen. Im Folgenden wird daher eine nah an der Erstellung der Variable MIGBACK angelegte Möglichkeit vorgestellt, vorhandene Informationen für eine möglichst breite Bestimmung der Herkunftsnationen der SOEP-Teilnehmer mit Migrationshintergrund zu verwenden. Dazu wird differenzierend der Beitrag einzelner Variablen bzw. Variablenblöcke zur Herkunftsbestimmung von Personen mit direktem, indirektem und zuvor nicht genauer bestimmbarem Migrationshintergrund dargestellt. Zudem werden Besonderheiten der SOEP-Teilnehmer mit bestimmbarer und nicht bestimmbarer Herkunft thematisiert. Weiterhin werden sich daraus eventuell ergebende Schwierigkeiten für weitere Analysen diskutiert und auf sonstige auftretende (u. a. Fallzahlen-)Probleme und Grenzen der skizzierten Vorgehensweise aufmerksam gemacht.

### **3        Herkunftsbestimmung: Vorgehensweise**

Um den SOEP-Teilnehmern ihre jeweiligen Herkunftsnationen zuordnen zu können, wurden die erwähnten Informationen zu einem vorhandenen Migrationshintergrund (MH) der Zugehörigen sämtlicher Stichproben<sup>2</sup> bis 2009 (kein MH / direkter MH / indirekter MH / MH, nicht weiter differenzierbar) zunächst direkt aus dem Datensatz übernommen (Variable MIGBACK). Der Beitrag einzelner Variablen und Variablenblöcke zur Herkunftsbestimmung bei sämtlichen jemals in der Bruttostichprobe enthaltenen Personen mit identifiziertem Migrationshintergrund kann nun Tabelle 1 entnommen werden. Bei der ersten Zuwanderergeneration stellt die Bestimmung der Herkunftsländer noch kaum ein Problem dar. Bei nahezu allen selbst Eingewanderten kann das Herkunftsland über die generierte Variable CORIGIN direkt bestimmt werden. So konnten mit diesem ersten Schritt 7.559 von 7.636 Personen mit direktem Migrationshintergrund die Herkunftsnationen zugeordnet werden, während bei nur 77 Personen die genaue Herkunft noch unklar bleibt (Wert in Klammern).

---

<sup>2</sup> Die neue Stichprobe I stellt hier noch eine Ausnahme dar, da diese Befragten in der ersten Befragungswelle den Biographiefragebogen noch nicht ausgefüllt haben und daher kein Migrationsstatus zugeordnet wurde (vgl. Groh-Samberg et al. 2010: 17).

Bei den Folgegenerationen, welche nicht mehr selbst eingewandert, sondern bereits in Deutschland geboren sind, gestaltet sich die Zuordnung schwieriger und ist nur bei einem niedrigeren Anteil überhaupt durchführbar. Die Variable CORIGIN kann keinen Beitrag zur Bestimmung der Herkunft der 4.779 Personen mit indirektem Migrationshintergrund leisten.

Tabelle 1

**Beitrag einzelner Variablen(-blöcke) zur Bestimmung der Herkunft (Bruttostichprobe)**

<i>Herkunft zusätzlich bestimmt bei ... Teilnehmern (in Klammern: noch fehlend / nicht bestimmt)</i>					
Schritt	Variable(n)	Datensatz/-sätze	direkter MH	indirekter MH	MH, n. w. d.
1	CORIGIN	ppfad	7.559 (77)	0 (4.779)	0 (1.142)
2	NATION\$\$ <sup>1</sup>	\$pgen	2 (75)	1.145 (3.634)	163 (979)
3	\$PNAT <sup>2</sup>	\$pbrutto	2 (73)	1.560 (2.074)	513 (466)
4	SP11702 <sup>3</sup>	sp	6 (67)	100 (1.974)	16 (450)
5	AK07A / EK03A <sup>4</sup>	\$kind	0 (67)	86 (1.888)	5 (445)
6	2. Staatsangeh. <sup>5</sup>	\$p	4 (63)	109 (1.779)	3 (442)
7	Geburtsland <sup>6</sup>	\$pausl/\$p/\$page17/\$pbrutto	7 (56)	122 (1.657)	7 (435)
8	Herkunft Eltern <sup>7</sup>	-	1 (55)	503 (1.154)	10 (425)
Anzahl insgesamt			7.581 (55)	3.625 (1.154)	717 (425)

MH = Migrationshintergrund, n. w. d. = nicht weiter differenzierbar

<sup>1</sup> Staatsangehörigkeit, generierte Variable: Wellen A-Z

<sup>2</sup> Staatsangehörigkeit nach Interviewerangaben („address log“): Wellen B-Z

<sup>3</sup> Frühere ausländische Staatsangehörigkeit, retrospektiv erfasst: Welle S

<sup>4</sup> Vom Haushaltsvorstand angegebene Staatsangehörigkeit eines Kindes im Haushalt: Wellen A und E

<sup>5</sup> Wellen Q-Z

<sup>6</sup> Diverse Quelldatensätze: Wellen A-L, Q, R, W, X, Y, Z

<sup>7</sup> Mittels der vorigen Schritte bestimmte Herkunft der Eltern, falls diese ebenfalls in der Stichprobe enthalten sind

Um soweit möglich eine Zuordnung der ethnischen Herkunft auch bei den zweiten und nachfolgenden Generationen zu erreichen, wurden in Schritt zwei daher zunächst eventuelle in der aktuellen oder in vergangenen Wellen erfasste nichtdeutsche Nationalitäten ermittelt (Variable NATION\$\$ aus \$pgen.dta<sup>3</sup>). Bei einigen Befragten ergibt sich hier das Problem, dass in verschiedenen Wellen unterschiedliche (ausländische) Nationalitäten angegeben wurden. In diesen Fällen wurde pragmatisch verfahren und die am häufigsten genannte Staatsangehörigkeit bzw. bei gleich häufiger Nennung die zuletzt genannte übernommen. Insgesamt führte die Berücksichtigung der Variable NATION\$\$ zur Zuordnung der Herkunftsnationen bei zusätzlich 1.145 Personen mit indirektem Migrationshintergrund, bei 163 Personen, bei denen die

<sup>3</sup> Die generierte Variable NATION\$\$ wird jährlich aus Angaben zur Staatsangehörigkeit erstellt, welche in den Datensätzen \$p und \$pausl enthalten sind (vgl. Anger et al. 2010: 38). Die entsprechenden Quellvariablen wurden teilweise ebenfalls bei der Bestimmung der Herkunft mit einbezogen, führen bei separater Berücksichtigung aber nicht zu einem zusätzlichen Informationsgewinn.

Art des Migrationshintergrunds „nicht weiter differenzierbar“ war sowie bei noch zwei Personen mit direktem Migrationshintergrund. Durch die beiden generierten Variablen CORIGIN und NATION\$\$ kann somit insgesamt bereits 8.869 Personen mit Migrationshintergrund ein Herkunftsland zugeordnet werden. Bei 3.634 Personen mit indirektem sowie 979 Personen mit nicht weiter differenzierbarem Migrationshintergrund und noch 75 derjenigen mit direktem Migrationshintergrund ist dies allerdings nicht möglich.

Bei denjenigen Personen, bei denen bislang kein Herkunftsland bestimmt werden konnte, wurden nun zunächst weitere Informationen zur Staatsangehörigkeit herangezogen (Schritte 3 bis 6 in Tabelle 1). Die verwendeten Variablen bzw. Variablenblöcke konnten dabei in unterschiedlichem Ausmaß zur zusätzlichen Herkunftsbestimmung bei SOEP-Teilnehmern mit Migrationshintergrund beitragen. Interviewerangaben zur Nationalität der Haushaltsmitglieder führen zur Herkunftsbestimmung in mehr als 2.000 Fällen (Schritt 3). Diese Informationen dienen insbesondere der Bestimmung der Herkunft bei Haushaltsmitgliedern, welche keine eigenen Angaben gemacht haben.<sup>4</sup> Bei auch hier auftretenden inkonsistenten Angaben wurde ebenso wie schon im zweiten Schritt verfahren. Die einmalig retrospektiv erhobene frühere Staatsangehörigkeit ermöglicht zusätzlich die Bestimmung bei weiteren 122 Personen, 100 hiervon mit indirektem Migrationshintergrund (Schritt 4). In zwei frühen Wellen gemachte Angaben des Haushaltsvorstands zur Staatsangehörigkeit von im Haushalt lebenden Kindern erlauben die Zuordnung bei etwa 90 Personen (Schritt 5). Weiterhin wurden der neuen Variable Angaben zu einer zweiten Staatsangehörigkeit, welche seit dem Jahr 2000 gemacht werden können, hinzugefügt (insgesamt +116 Fälle, Schritt 6).

Bei 63 Personen mit direktem, 1.779 Personen mit indirektem und 442 Personen mit nicht weiter differenzierbarem Migrationshintergrund ist die Erfassung anhand aktueller bzw. früherer Staatsangehörigkeiten weiterhin nicht möglich. In diesen Fällen wurden zusätzlich zu den zusammengefassten Informationen der generierten Variable CORIGIN noch einmal angegebene Geburtsländer der Befragten aus verschiedenen Datensätzen zusammengetragen und diese als Herkunftsländer identifiziert (weitere 136 Personen, Schritt 7 in Tabelle 1). Wie bereits zuvor sind sowohl bei der zweiten Staatsangehörigkeit als auch beim Geburtsland

---

<sup>4</sup> In der Regel sind in den verwendeten Variablen sämtliche Herkunftsnationen differenziert enthalten. Bei den Interviewerangaben zur Staatsangehörigkeit in den Variablen \$PNAT sowie teilweise auch in den Daten zum Geburtsland werden in früheren Erhebungsjahren jedoch nur die Nationalitäten der Ausländerstichprobe B (Türkei, (Ex-)Jugoslawien, Griechenland, Italien und Spanien) separat aufgeführt, während andere Nationen unter „Sonstige“ zusammengefasst sind. Daher ist es möglich, dass bei separater Betrachtung einzelner Staatsangehörigkeiten die Anzahl der Migranten mit bestimmter Herkunft etwas niedriger ist, als hier dargestellt.

in einigen Fällen zu verschiedenen Befragungszeitpunkten inkonsistente Angaben festzustellen. Hier wurde die jeweils zuletzt gemachte Angabe als ‚wahr‘ eingeschätzt.

Bei 122 Personen lässt sich (bei der vorgestellten hierarchischen Vorgehensweise) das Herkunftsland über ein in mindestens einer Welle angegebenes Geburtsland außerhalb Deutschlands bestimmen, obwohl diese nach der Zuordnung in der Variable MIGBACK lediglich einen indirekten Migrationshintergrund haben. Große Anteile entfallen hierbei auf die Herkunftsländer Tschechien (53 Personen) und Polen (24). Nur in wenigen Fällen stammen die Personen aus den ehemaligen Anwerbeländern. Da nur Zuwanderer, die nach 1949 nach Deutschland immigriert sind als Personen mit direktem Migrationshintergrund behandelt werden, ist der Grund für dieses zunächst überraschende Ergebnis vor allem in einem frühen Zuwanderungsdatum der betreffenden Personen zu sehen.<sup>5</sup> Allerdings wurden immerhin 21 dieser Personen (davon sechs aus den ehemaligen Anwerbestaaten) erst nach 1949 geboren und können damit erst zu einem späteren Zeitpunkt zugewandert sein, so dass dies nicht die einzige Erklärung ist.<sup>6</sup> In jedem Fall tragen auch die hier gewonnenen Daten zur besseren Darstellung der Heterogenität der Population mit Migrationshintergrund bei, weshalb diese in die erstellte Variable mit aufgenommen werden.

Über die direkten Angaben der SOEP-Teilnehmer hinaus können abschließend zusätzlich die Vorteile einer Haushaltsbefragung genutzt werden. So können den Befragten über die Zeiger MNR und VNR im Datensatz BIOPAREN Merkmale ebenfalls in der Stichprobe enthaltener Eltern zugeordnet werden (Schritt 8). Bei unterschiedlichen Herkunftsnationen wurde dem Kind die Herkunft der Mutter übertragen. Ist die Mutter deutscher, der Vater aber ausländischer Herkunft, wurde die Herkunft des Vaters übernommen. Auf diesem Wege wurden bei 503 Personen mit indirektem Migrationshintergrund, zehn ‚nicht differenzierbaren‘ und noch einer selbst eingewanderten Person die zuvor auf die vorgestellte Weise ermittelten Herkunftsnationen der Eltern übernommen (insgesamt 514 Fälle).

Im Ergebnis kann 7.581 Personen mit direktem Migrationshintergrund (nicht bestimmt: 55), 3.625 Personen mit indirektem Migrationshintergrund (n. b.: 1.154) und noch 717 Personen, deren Migrationshintergrund nicht genauer bestimmbar ist (n. b.: 425), ein Herkunftsland zugeordnet werden. Allerdings beziehen sich diese Zahlen auf sämtliche Zugehörigen der

---

<sup>5</sup> Vermutlich handelt es sich hier v. a. um Vertriebene bzw. „ethnische Deutsche“. Die Bedeutung dieser Personengruppe ist letztlich abhängig von den Inhalten geplanter Untersuchungen. Beispielsweise sollten diese Personen für Analysen des Arbeitsmarktes aufgrund des eher hohen Alters eine untergeordnete Rolle spielen.

Bruttostichprobe, welche auch diejenigen Personen enthält, welche nie tatsächlich an der Befragung teilgenommen haben. Tabelle 2 bildet daher die einzelnen Bestimmungsschritte nur für Personen ab, für die mindestens einmal auch tatsächlich persönliche Angaben verfügbar sind, die über allgemeine Haushaltsinformationen hinausgehen. Neben der Teilnahme über ein eigenes Interview zählen hierzu zum Beispiel bei Kindern auch Proxy-Angaben der Eltern im Mutter-Kind-Fragebogen oder ein ausgefüllter Jugendfragebogen. Vor allem die Anzahl der Personen mit indirektem (jetzt insgesamt 2.633) oder nicht weiter differenzierbarem Migrationshintergrund (427) reduziert sich hierdurch deutlich.

Tabelle 2

**Beitrag einzelner Variablen(-blöcke) zur Bestimmung der Herkunft (nur Personen mit mind. einmaligen persönlichen Angaben)**

<i>Herkunft zusätzlich bestimmt bei ... Teilnehmern (in Klammern: noch fehlend / nicht bestimmt)</i>					
Schritt	Variable(n)	Datensatz/-sätze	direkter MH	indirekter MH	MH, n. w. d.
1	CORIGIN	ppfad	7.268 (71)	0 (2.633)	0 (427)
2	NATION\$ <sup>1</sup>	\$pgen	2 (69)	1.145 (1.488)	163 (264)
3	\$PNAT <sup>2</sup>	\$pbrutto	2 (67)	176 (1.312)	269 (197)
4	SP11702 <sup>3</sup>	sp	6 (61)	100 (1.212)	16 (181)
5	AK07A / EK03A <sup>4</sup>	\$kind	0 (61)	14 (1.198)	0 (181)
6	2. Staatsangeh. <sup>5</sup>	\$p	4 (57)	109 (1.089)	3 (178)
7	Geburtsland <sup>6</sup>	\$pausl/\$p/\$page17/\$pbrutto	7 (50)	115 (974)	0 (178)
8	Herkunft Eltern <sup>7</sup>	-	1 (49)	503 (471)	10 (168)
Anzahl insgesamt			7.290 (49)	2.162 (471)	259 (168)

MH = Migrationshintergrund, n. w. d. = nicht weiter differenzierbar

<sup>1</sup> Staatsangehörigkeit: Wellen A-Z

<sup>2</sup> Staatsangehörigkeit nach Interviewerangaben („address log“): Wellen B-Z

<sup>3</sup> Frühere ausländische Staatsangehörigkeit, retrospektiv erfasst: Welle S

<sup>4</sup> Vom Haushaltsvorstand angegebene Staatsangehörigkeit eines Kindes im Haushalt: Wellen A und E

<sup>5</sup> Zweite Staatsangehörigkeit: Wellen Q-Z

<sup>6</sup> Geburtsland, diverse Quelldatensätze: Wellen A-L, Q, R, W, X, Y, Z

<sup>7</sup> Mittels der vorigen Schritte bestimmte Herkunft der Eltern, falls diese ebenfalls am SOEP teilgenommen haben

Es zeigt sich, dass die zusätzliche Bestimmung der Herkunft durch die Interviewerinformationen (\$PNAT) in der Bruttostichprobe weitestgehend bei Personen stattfindet, die zwar Teil von Befragungshaushalten sind bzw. waren, aber niemals selbst an der Befragung teilgenommen haben. Dennoch trägt diese Variable immer noch zur Bestimmung der Herkunft von 176

<sup>6</sup> Für sämtliche betroffenen Personen nimmt die Variable IMMIYEAR den Wert -2 an („born in Germany or immigrated before 1949, respectively and survey participants without an interview“). Zuwanderungsjahre vor 1949 sind kaum bestimmbar, daher wurde das Geburtsjahr als Indikator verwendet.

Personen mit indirektem und 269 Personen mit nicht genau bestimmbarem Migrationshintergrund bei, für die persönliche Angaben vorhanden sind. Geringer ist nun auch der Ertrag der Elternangaben zu eigenen Kindern in frühen Erhebungswellen (Schritt 5). Diese können nur noch in 14 Fällen einen positiven Beitrag leisten. Auch die Angaben zum Geburtsland (Schritt 7) entstammen zum Teil den Daten der Bruttostichprobe und können daher nun etwas weniger häufig für die Herkunftsbestimmung genutzt werden. Die zuvor über die letztgenannten Schritte zusätzlich zugeordneten Personen haben letztlich in keiner Form persönlich an der Befragung teilgenommen.

#### 4 Migrantengruppen mit bestimmter / nicht bestimmter Herkunft

Wie gezeigt stellt die Bestimmung der Herkunftsnation bei nahezu sämtlichen selbst Einwanderten kein Problem dar. Betrachtet man nur die für die meisten Analysen relevante Gruppe der Personen, für die persönliche Informationen verfügbar sind, ist für 7.268 von 7.339 Personen mit direktem Migrationshintergrund die benötigte Information bereits in der Variable CORIGIN enthalten und musste lediglich übernommen werden. Während die Herkunft damit bei 99% der Mitglieder der ersten Generation problemlos bestimmt werden kann, konnten weitere Variablen zumindest bei einem Großteil derjenigen mit indirektem Migrationshintergrund zur Herkunftsbestimmung beitragen.

Tabelle 3  
Fallzahlen / Anteile Herkunftsbestimmung

		Herkunft bestimmt	Herkunft n. bestimmt	N
Gesamt (Brutto)		11.924 (88,0%)	1.633 (12,0%)	13.557
<i>Davon:</i>				
mind. 1x mit persönlichen Angaben	direkter MH	7.290 (99,3%)	49 (0,7%)	7.339
	indirekter MH	2.162 (82,1%)	471 (17,9%)	2.633
	MH, n. w. d.	259 (60,7%)	168 (39,3%)	427
	Gesamt	9.711 (93,4%)	688 (6,6%)	10.399
mindestens einmal ...	im Alter 17-64 (Brutto)	10.967 (94,9%)	595 (5,1%)	11.562
	mit erfolgreichem Personeninterview <sup>1</sup>	9.503 (96,2%)	372 (3,8%)	9.875

MH = Migrationshintergrund, n. w. d. = nicht weiter differenzierbar, LFS = Labour Force Status

<sup>1</sup> ausgefüllter Erwachsenenfragebogen (ohne Jugendfragebögen, Mutter-Kind-Interviews etc.)

So wurde bei 2.162 Personen mit indirektem Migrationshintergrund (82%, fehlend: 471) und bei 259 Personen mit zuvor nicht genauer bestimmtem Migrationshintergrund (61%, fehlend: 168) das Herkunftsland identifiziert. Insgesamt konnte die Herkunft von 93,4% der Migranten mit mindestens einmaligen persönlichen Angaben bestimmt werden (siehe Tabelle 3 oben). Betrachtet man lediglich die Gruppe, bei der tatsächlich mindestens einmal ein SOEP-Standardinterview realisiert werden konnte (ohne Berücksichtigung von Mutter-Kind-Interviews etc.), erhöht sich der Anteil derjenigen, deren Herkunft bestimmt ist sogar auf 96% (unterer Teil von Tabelle 3). Bei Betrachtung der gesamten Bruttostichprobe beträgt dieser Anteil hingegen nur 88 Prozent.<sup>7</sup>

Trotz der insgesamt hohen Anteile von Migranten im SOEP, denen ein Herkunftsland zugeordnet werden konnte, ist die erste Zuwanderergeneration hier überrepräsentiert. Während die Daten von Migranten der ersten Generation zu nahezu 100% für herkunftsspezifische Analysen verwendet werden können, trifft dies nur für vier Fünftel der generell schon in geringerer Zahl vertretenen Migranten der Folgegenerationen zu, setzt man vorhandene persönliche Informationen voraus. Dieses Problem lässt sich bei der momentanen Datenlage nicht lösen und könnte bei Nichtberücksichtigung der Migranten ohne bestimmte Herkunftsnation zu verzerrten Ergebnissen führen. Beispielhaft wurden daher für Personen mit indirektem Migrationshintergrund im Jahr 2009 Unterschiede zwischen denjenigen mit bestimmbarer Herkunft und denjenigen, deren Herkunft nicht zuzuordnen war, untersucht.

Ein Vergleich des durchschnittlichen Alters weist bereits darauf hin, dass Unterschiede hier nicht rein zufällig sind. Während diejenigen Personen mit indirektem Migrationshintergrund und persönlichen Angaben, deren Herkunft bestimmt werden konnte, im Jahr 2009 im Durchschnitt 34,7 Jahre alt sind (N=789), beträgt das Durchschnittsalter der 233 Migranten nachfolgender Generationen ohne bestimmte Herkunft in diesem Erhebungsjahr 41,4 Jahre. Mittels logistischer Regressionsmodelle wurde nun der Frage nachgegangen, ob entlang weiterer Variablen signifikante Unterschiede zwischen beiden Gruppen bestehen (Tabelle 4). Berücksichtigt wurden Personen ab 18 Jahren mit indirektem Migrationshintergrund und Angaben zur Berufsausbildung im Jahr 2009. Bei alleiniger Betrachtung der unabhängigen Variablen Alter und Geschlecht werden zunächst die bereits beschriebenen Altersunterschiede deutlich. Je älter eine Person mit indirektem Migrationshintergrund im SOEP ist, desto seltener kann im Jahr 2009 ihre Herkunft bestimmt werden. Keine signifikanten Unterschiede zeigen sich

---

<sup>7</sup> Da für die meisten Untersuchungen nur Personen relevant sind, für die individuelle Informationen vorliegen, wird im Folgenden ausschließlich diese Gruppe betrachtet.

hingegen bezüglich des Geschlechts (Modell 1, abhängige Variable: Herkunft bestimmt). Bei zusätzlicher Kontrolle des höchsten beruflichen Bildungsabschlusses (Modell 2) bleibt der signifikante Alterseffekt bestehen. Zusätzlich zeigt sich, dass bei Personen mit erfolgreicher Berufsausbildung signifikant seltener die Herkunft bestimmt werden kann, als bei Personen ohne beruflichen Abschluss.

Tabelle 4

**Logistische-Regression, abhängige Variable Herkunft bestimmt vs. Herkunft nicht bestimmt, Personen ab 18 Jahren mit indirektem Migrationshintergrund und persönlichen Angaben im Jahr 2009**

Variable	Modell 1	Modell 2	Modell 3
Alter	-0,03*** (0,00)	-0,03*** (0,00)	-0,02** (0,00)
Geschlecht: männlich	-0,04 (0,17)	0,02 (0,17)	0,08 (0,19)
<i>Beruflicher Abschluss (Referenz: kein Abschluss)</i>			
In Ausbildung	-	-0,49 (0,37)	-0,25 (0,40)
Ausbildungsabschluss	-	-0,78** (0,24)	-0,47 (0,26)
Hochschulabschluss	-	-0,99*** (0,28)	-0,64* (0,31)
<i>Stichprobe (Referenz: Stichprobe A)</i>			
B	-	-	2,37*** (0,45)
C	-	-	0,10 (0,44)
D	-	-	-0,02 (0,60)
E	-	-	-0,46 (0,47)
F	-	-	0,04 (0,24)
G	-	-	-0,12 (0,36)
H	-	-	-2,31*** (0,34)
McFadden's R <sup>2</sup>	0,06	0,08	0,24
N	917	917	917

Standardfehler in Klammern; \*\*\* p<0,001, \*\* p<0,01, \* p<0,05

Dass dieser Bildungseffekt im Kern nicht auf ein ablehnenderes Antwortverhalten höher Gebildeter zurückzuführen ist, lässt Modell 3 vermuten. Berücksichtigt man die Stichprobenzugehörigkeit der Teilnehmer, ist bei einem nun deutlich größeren Anteil erklärter Varianz nur noch ein leicht signifikanter negativer Zusammenhang zwischen der Bestimmung der Herkunft und einem vorhandenen Hochschulabschluss zu finden. Die in einigen Fällen deutlichen Unterschiede zwischen verschiedenen Stichproben überraschen dabei nicht, setzen sich diese doch bezüglich soziodemografischer Merkmale wie auch der Migrantenanteile teils sehr unterschiedlich zusammen. Zudem variiert die Dauer der Teilnahme an der Befragung bei den entsprechenden Haushalten stark. Im Vergleich zu Personen mit indirektem Migrationshinter-



grund aus Haushalten der Stichprobe A aus dem Jahr 1984 kann entsprechend besonders häufig die Herkunft von Personen aus Haushalten der Ausländerstichprobe B bestimmt werden. Da diese sich aus Migranten der ehemaligen Gastarbeiteranwerbestaaten mit durchschnittlich eher niedrigem Bildungsniveau zusammensetzt, ist anzunehmen, dass der zuvor beobachtete negative Bildungseffekt hauptsächlich auf die bessere Bestimmbarkeit der Herkunft bei diesen Personen zurückzuführen ist. Die Bedeutung der Stichprobenzugehörigkeit wurde damit nur überdeckt. Deutlich negativ wirkt sich hingegen die Zugehörigkeit zur Ergänzungsstichprobe 2006 aus. Hier kann signifikant seltener die Herkunft bestimmt werden als bei Personen mit indirektem Migrationshintergrund aus Stichprobe A. Aufgrund der erst kurzen Teilnahme an der Befragung erscheint es plausibel, dass bei Haushalten der Stichprobe H weniger Herkunftsinformationen existieren, so dass hier z. B. die Übernahme von Informationen der Eltern seltener möglich ist.

Das Beispiel zeigt, dass im Jahr 2009 vor allem bei Personen aus Haushalten der Ergänzungsstichprobe H signifikant seltener die Herkunft bestimmt werden kann, während dies bei Haushaltsmitgliedern der Ausländerstichprobe B auch in den Folgegenerationen häufiger möglich ist. Jedoch bleiben auch bei Berücksichtigung der Stichprobenzugehörigkeit sowohl ein leichter Ausbildungseffekt als auch ein signifikanter Einfluss des Alters bestehen. Neben diesen Faktoren können zudem weitere Merkmale existieren, in denen sich die Gruppen mit bestimmter und nicht bestimmter Herkunft unterscheiden und welche über spezifische Stichprobenmerkmale hinausgehen. Es ist weiterhin zu beachten, dass es sich hier nur um ein Beispiel der Zusammensetzung in der letzten verfügbaren Erhebungswelle 2009 handelt. In anderen Erhebungsjahren können die Zusammensetzung beider Gruppen und damit die relevanten Merkmale andere sein. Mögliche Verzerrungen, welche durch die unvollständige Herkunftszuordnung bedingt sein könnten, sollten daher immer bedacht werden.

## **5 Verteilung nach Herkunftsländern**

Eine nach den identifizierten Herkunftsländern differenzierte Betrachtung der SOEP-Teilnehmer spiegelt die zu Beginn starke Fokussierung der Erhebung auf Migranten aus den ehemaligen Gastarbeiteranwerbestaaten wider. Im Folgenden sollen beispielhaft insbesondere diese Herkunftsländer betrachtet werden, wobei sich die Darstellung auf die Unterschiede der Herkunftsbestimmung bei ausschließlicher Berücksichtigung der generierten Variable NATION\$\$ gegenüber der Durchführung der hier vorgestellten zusätzlichen Schritte bezieht. Da, wie gezeigt, bei Migranten der ersten Generation die Herkunft in der Regel bereits anhand der

Variable CORIGIN bestimmbar ist, werden diese Differenzen für Personen mit indirektem Migrationshintergrund und vorhandenen individuellen Informationen dargestellt (Tabelle 5).

Bei ausschließlicher Bestimmung der Herkunft von Zugehörigen der zweiten bzw. dritten Zuwanderergeneration über die Variablen CORIGIN und NATION\$\$ stammen über 90% der betreffenden Personen aus der Türkei, Ex-Jugoslawien, Griechenland, Italien, Spanien oder Portugal. Nur etwa jeder Zwanzigste entstammt sonstigen Herkunftsländern. Mit knapp 37% der identifizierten Personen mit Migrationshintergrund stellen SOEP-Teilnehmer mit türkischem Hintergrund über alle Befragungswellen seit 1984 die mit Abstand größte Einzelgruppe unter den identifizierten Migranten der Nachfolgegenerationen dar. 422 Befragte sind dieser Gruppe zuzuordnen. Größere Fallzahlen finden sich weiterhin bei Teilnehmern (ex-)jugoslawischer, griechischer und italienischer Herkunft. Befragte aus den weiteren Anwerbeländern Spanien und Portugal sind entsprechend ihrem Anteil in der Gesamtbevölkerung bereits in geringerer Anzahl vertreten.

Tabelle 5

**Herkunftsnationen der SOEP-Teilnehmer mit indirektem Migrationshintergrund (nur Personen mit mind. einmaligen persönlichen Angaben)**

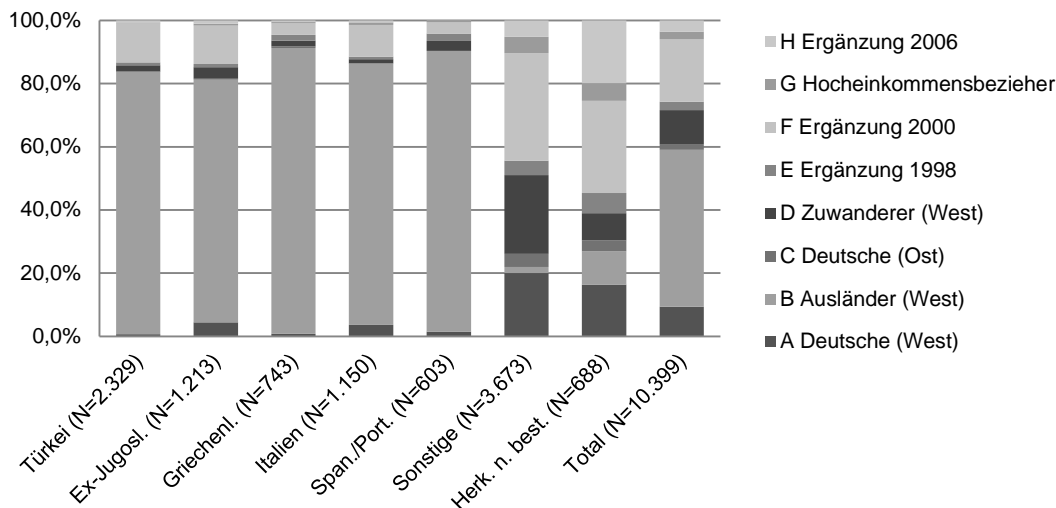
Herkunftsland	<i>Nur mit Berücksichtigung von NATION\$\$</i>		<i>Bei Berücksichtigung der weiteren Schritte</i>		
	N	Anteil	zusätzlich bestimmt	N	Anteil
Türkei	422	36,9%	+105	527	24,3%
Ex-Jugoslawien	171	14,9%	+65	236	10,9%
Griechenland	155	13,5%	+28	183	8,5%
Italien	236	20,6%	+110	346	16,0%
Spanien/Portugal	98	8,6%	+36	134	6,2%
Sonstige	63	5,5%	+673	736	34,0%
Gesamt	1.145	100,0%	+1.017	2.162	100,0%

Einem relativ großen Anteil der Personen mit indirektem Migrationshintergrund kann jedoch anhand der ersten beiden Schritte kein Herkunftsland zugeordnet werden. Werden nun zusätzlich die weiteren Schritte der Herkunftsbestimmung durchgeführt, ändert sich das Bild deutlich (Tabelle 5 rechts). Durch diese Schritte kann bei 673 Personen mit einem ‚sonstigen‘ indirekten Migrationshintergrund ein Herkunftsland bestimmt werden, so dass sich der Anteil dieser Gruppe von etwa fünf Prozent auf 34 Prozent steigert. Betroffen sind hier insbesondere

Personen osteuropäischer Herkunft mit deutscher Staatsangehörigkeit. Trotz des nun geringeren Anteils können aber auch bei den ehemaligen Anwerbestaaten nun höhere Zahlen von Personen mit bestimmter Herkunft verzeichnet werden. Dies betrifft mit zusätzlich bestimmter Herkunft in jeweils mehr als 100 Fällen besonders die generell in größerer Zahl vertretenen Herkunftsländer Türkei und Italien, kann aber auch bei den weiteren ehemaligen Anwerbeländern beobachtet werden. Die insgesamt 1.017 durch die zusätzlichen Schritte drei bis acht (vgl. Tabellen 1 und 2) neu hinzugewonnenen Fälle steigern den Anteil der potentiell auf Herkunftsunterschiede untersuchbaren SOEP-Teilnehmer damit auch für die hier differenzierter betrachteten Herkunftsländer.

Abbildung 1

**SOEP-Teilnehmer mit Migrationshintergrund und mindestens einmaligen persönlichen Angaben, nach Stichproben**

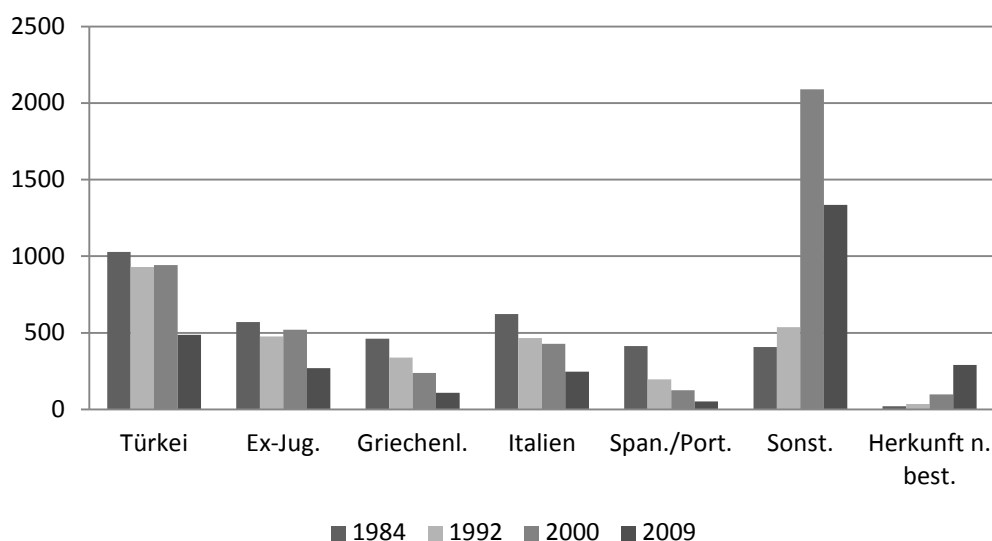


Obwohl die zusätzlichen Bestimmungsschritte den Anteil der Personen mit sonstiger Herkunft erhöhen, beträgt der Anteil der Personen mit indirekter Herkunft aus den ehemaligen Anwerbestaaten weiterhin über drei Fünftel. Die frühe Fokussierung auf Migranten der ehemaligen Anwerbestaaten zeigt sich erneut bei Betrachtung der Verteilung der Gesamtzahl der Migranten, unabhängig von der Art des Migrationshintergrundes, auf die verschiedenen SOEP-Stichproben (Abbildung 1): Teilnehmer mit einer Herkunft aus den ehemaligen Anwerbestaaten sind zu ca. 80% der Ausländerstichprobe B aus dem Jahr 1984 zuzuordnen. Wie zuvor beobachtet ist in dieser Stichprobe die Zuordnung häufiger möglich (vgl. Tabelle 4). Personen mit sonstiger ausländischer Herkunft sind hingegen vor allem in der Zuwandererstichprobe D (1994/1995) und in der Ergänzungsstichprobe F (2000) zu finden und damit erst später in die Erhebung aufgenommen worden. Weiterhin befinden sich SOEP-Teilnehmer mit Migrations-

hintergrund aus Stichprobe H überproportional häufig in der Gruppe ohne bestimmte Herkunft.

Entsprechend stellt sich auch die Verteilung der Herkunftsgruppen auf die verschiedenen Erhebungswellen dar. Während in der Gesamtbetrachtung die Zahl der SOEP-Teilnehmer aus den Anwerbestaaten die Zahl derjenigen aus sonstigen Ländern deutlich übersteigt, ändert sich dieses Verhältnis bei separater Betrachtung einzelner Wellen. Betrachtet man die Entwicklung der nach Herkunftsländern differenzierten Fallzahlen über die Zeit, so zeigt sich unter den Teilnehmern aus den Anwerbestaaten ein eindeutiger negativer Trend (Abbildung 2). Die im Acht- bzw. Neun-Jahres-Abstand dargestellten Werte zeigen eine kontinuierliche Abnahme der Teilnehmerzahlen vor allem bei den generell bereits in geringerer Zahl vertretenen Herkunftsnationen. In der letzten hier betrachteten Befragungswelle im Jahr 2009 sinkt die Teilnahme bei den ‚traditionellen‘ Herkunftsländern auf höchstens knapp unter die Hälfte (Türkei, Ex-Jugoslawien) bis nur noch etwa 13% (Spanien/Portugal, 2009: 52 Teilnehmer) der ursprünglichen Teilnehmerzahlen. Dieser Rückgang ist zum einen durch das Alter der Ausländerstichprobe und die damit verbundene Panelmortalität bedingt. Zum anderen sollten hier aber auch Unterschiede zwischen den Herkunftsländern hinsichtlich der Rückwanderung und der Haushaltsgröße bzw. der Anzahl der Mitglieder der Folgegenerationen eine Rolle spielen. Bei der ex-jugoslawischen Herkunftsgruppe ist zudem die zahlenmäßige Bedeutung der zweiten Zuwanderungswelle in Folge der kriegesischen Auseinandersetzungen zu berücksichtigen, womit die hier zwischenzeitlich wieder ansteigenden Teilnehmerzahlen erklärt werden könnten.

Abbildung 2  
**Entwicklung der Fallzahlen im Zeitverlauf, nach Herkunftsländern**



Im Gegensatz zu den bei den bisher genannten Herkunftsnationen stetig fallenden Teilnehmerzahlen verläuft die Entwicklung der Teilnahme der Gruppe der ‚sonstigen‘ Migranten, welche sich zu einem Großteil aus Menschen osteuropäischer bzw. russischer Herkunft zusammensetzt, entgegengesetzt. Ist die Zahl der Sonstigen mit 407 Befragten im Erstbefragungsjahr 1984 noch relativ gering, so steigert sie sich bis auf über 2.000 Befragte zur Jahrtausendwende und liegt auch im Jahr 2009 im Gegensatz zur ersten Befragungswelle über der Gesamtzahl der Befragten mit Herkunft aus den ausgewählten Anwerbeländern. Der dargestellte Verlauf ist insbesondere der Zuwanderung von Spätaussiedlern und deren Aufnahme in die Befragung anhand neuer Stichproben geschuldet und spiegelt im Ergebnis die heute heterogenere Zusammensetzung der zugewanderten Bevölkerung in Deutschland wider.

## **6        Schluss**

Es konnte gezeigt werden, dass es die Daten des Sozio-ökonomischen Panels für einen Großteil der dort identifizierten Migranten bzw. Personen mit Migrationshintergrund ermöglichen, neben dem Migrationsstatus auch nachträglich für frühere SOEP-Teilnehmer die jeweilige Herkunftsnation zu bestimmen. Dies gilt nicht nur für selbst Eingewanderte und für Menschen ausländischer Nationalität, sondern ist anhand verschiedener Variablen auch für Angehörige der Folgegenerationen mit deutscher Staatsangehörigkeit möglich. Zwar kann die Zuordnung nicht bei sämtlichen Personen mit Migrationshintergrund erfolgen, jedoch wurde deutlich, dass fehlende Informationen zum Herkunftsland häufig mit generell fehlenden individuellen Informationen einhergehen, so dass dieser Personenkreis für die meisten Analysen nur eine untergeordnete Rolle spielen sollte. Das Problem unbestimmbarer Herkunftsnationen sollte sich damit kaum in einem Verlust einer größeren Anzahl von Fällen auswirken.

Entsprechend der in der ersten Erhebungswelle expliziten Konzentration auf Migranten aus den ehemaligen Anwerbestaaten und deren damit verbundener Überrepräsentierung entstammt die Mehrheit der jemals am SOEP beteiligten Migranten diesen Herkunftsgruppen. Auch in der Gesamtbetrachtung stellt bei diesen Gruppen die ‚eigene‘ bereits im Jahr 1984 befragte Stichprobe B den mit Abstand größten Anteil. Weniger als ein Fünftel der SOEP-Teilnehmer mit Herkunft aus den Anwerbestaaten entstammt einer der anderen Stichproben. Anders gestaltet sich die Verteilung bei der Betrachtung weiterer Herkunftsstaaten. Hier ist die Verteilung auf einzelne Stichproben insgesamt weitaus heterogener. Dies wirkt sich auch auf die Zusammensetzung der Bevölkerung mit Migrationshintergrund im SOEP in einzelnen Jahren aus. Hier sind im Zeitverlauf deutliche Schwankungen und gegenläufige Tendenzen zu

erkennen. Die ‚Dominanz‘ der ehemaligen Anwerbestaaten wird dabei entsprechend der demografischen Realität weitgehend aufgelöst.

Vor allem für migrationsbezogene Analysen, welche längere Zeiträume betrachten, muss die besondere Bedeutung der variierenden Zusammensetzung und tendenziell steigenden Heterogenität der Bevölkerung mit Migrationshintergrund hervorgehoben werden. Neben der hinsichtlich ihrer Herkunftsländer deutlich veränderten Zusammensetzung der Bevölkerung mit Migrationshintergrund ist zudem der heute größere Anteil von Migranten der zweiten und späteren Generationen zu berücksichtigen. Mit den für einzelne Herkunftsnationen sehr deutlich fallenden Teilnehmerzahlen geht ein sinkender Grad an Repräsentativität der entsprechenden Daten einher, welcher herkunftsnationenspezifische Analysen über längere Zeitabschnitte deutlich erschwert. Die zweite und dritte Migrantengeneration sind hiervon besonders betroffen. Auch mittels Gewichtung kann dieses Problem kaum gelöst werden (vgl. Diehl/Schnell 2006: 798). Bei einem Teil der Personen mit indirektem oder nicht weiter differenzierbarem Migrationshintergrund kann die Herkunft weiterhin nicht bestimmt werden. Soweit dies nicht in einem jungen Alter und daher noch fehlenden (Lebenslauf-)Informationen dieser Befragten begründet ist, könnte hier evtl. bereits eine retrospektive Nacherhebung helfen. Weiterhin wäre hier möglicherweise an eine erneute Auffrischung der Stichprobe mit besonderer Berücksichtigung der größten Herkunftsnationen zu denken, um der Panelmortalität entgegenzuwirken. Dies insbesondere auch mit dem Ziel, Analysen der zu Erhebungsbeginn im Jahr 1984 überrepräsentierten Herkunftsnationen auch zukünftig anhand vertretbarer Fallzahlen zu ermöglichen.

## Literatur

- Anger, Silke et al. / Projektgruppe SOEP (2010): Documentation PGEN. Person-related status and generated variables. Berlin. (Online verfügbar unter [http://www.diw.de/documents/dokumentenarchiv/17/diw\\_01.c.60055.de/pgen.pdf](http://www.diw.de/documents/dokumentenarchiv/17/diw_01.c.60055.de/pgen.pdf))
- Berger, Peter A. und Anja Weiß (Hg.) (2008): Transnationalisierung sozialer Ungleichheit. Wiesbaden.
- Diefenbach, Heike und Anja Weiß (2006): Menschen mit Migrationshintergrund. Datenerfassung für die Integrationsberichterstattung. Gutachten für das Statistische Amt & die Stelle für interkulturelle Arbeit der Landeshauptstadt München. München.
- Diehl, Claudia und Rainer Schnell (2006): "Reactive Ethnicity" or "Assimilation"? Statements, Arguments, and First Empirical Evidence for Labor Migrants in Germany. *International Migration Review* 40 (4): 786–816.
- Granato, Nadia (2009): Effekte der Gruppengröße auf die Arbeitsmarktintegration von Migranten. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 61 (3): 387-409.
- Gresch, Cornelia und Cornelia Kristen (2011): Staatsbürgerschaft oder Migrationshintergrund? Ein Vergleich unterschiedlicher Operationalisierungsweisen am Beispiel der Bildungsbeteiligung. *Zeitschrift für Soziologie* 40 (3): 208-227.
- Groh-Samberg, Olaf, Florian R. Hertel und Ingrid Tucci (2010): Variables MIGBACK and MIGINFO. S. 17-19 in: Frick, Joachim R. und Henning Lohmann (Hg.): *Biography and Life History Data in the German Socio Economic Panel (SOEP, v26, 1984-2009)*. Berlin. (Online verfügbar unter [http://www.diw.de/documents/publikationen/73/diw\\_01.c.360885.de/diw\\_datadoc\\_2010-052.pdf](http://www.diw.de/documents/publikationen/73/diw_01.c.360885.de/diw_datadoc_2010-052.pdf))
- Frick, Joachim R. und Janina Söhn (2005): Das Sozio-oekonomische Panel (SOEP) als Grundlage für Analysen zur Bildungslage von Personen mit Migrationshintergrund. S. 81-90 in: BMBF (Hg.): *Migrationshintergrund von Kindern und Jugendlichen. Wege zur Weiterentwicklung der amtlichen Statistik. Bildungsreform Band 14*. Bonn/Berlin.
- Haug, Sonja (2005). Interethnische Kontakte, Homogenität und Multikulturalität der Freundesnetzwerke. S. 251-275 in: Haug, Sonja und Claudia Diehl (Hg.): *Aspekte der Integration. Eingliederungsmuster und Lebenssituation italienisch- und türkischstämmiger junger Erwachsener in Deutschland*. Schriftenreihe des Bundesinstituts für Bevölkerungsforschung, Bd. 35. Wiesbaden.
- Haug, Sonja (2011): Migration and Statistics. In: Rat für Sozial- und Wirtschaftsdaten (Hg.): *Building on Progress. Expanding the Research Infrastructure for the Social, Economic, and Behavioral Sciences*. Vol. 2. Opladen.
- Kalter, Frank (2007): Ethnische Kapitalien und der Arbeitsmarkterfolg Jugendlicher türkischer Herkunft. S. 393-417 in: Wohlrab-Sahr, M. und L. Teczan (Hg.): *Konfliktfeld Islam in Europa. Soziale Welt, Sonderband*. Baden-Baden.
- Pries, Ludger (2008): *Die Transnationalisierung der sozialen Welt. Sozialräume jenseits von Nationalgesellschaften*. Frankfurt am Main.